②

USING THE MEDICAL INFORMATION PROJECTION SYSTEM TO ORGANIZE AND PRESENT BATTLE CASUALTY DATA

AD-A239 775

R. RIEMER

DTIC
ELECTE
S AUG 2 2 1991
B D

20030211045

REPORT No. 91-6

91-08374

91 8 20 119

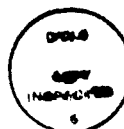Using the Medical Information Projection System
to Organize and Present Battle Casualty Data

Richard Riemer

Naval Health Research Center
Medical Decisions Support Department
P.O. Box 85122
San Diego, Ca 92186-5122

TABLE OF CONTENTS

SUMMARY

Problem

Medical resource planning requires projections of the number of casualties expected during military operations. Though there is a considerable amount of archival data on casualties, this information is not generally in a format compatible with casualty rate projections.

Objective

The objective of this effort is to show how battle casualty (BC) data from WWII can be adapted to the Medical Information Projection System (MIPS). In addition this document will present the features of MIPS that enhance the manipulation and evaluation of data.

Approach

This report will detail the use of MIPS in preparing the BC data and then evaluate the benefits and limitations of using MIPS with the BC data. First, the statistical principles used in MIPS will be explained. Next the process of analyzing raw data and determining what information will be needed to configure MIPS is demonstrated. The sequence of MIPS commands needed to configure MIPS will be specified. After MIPS has been configured for the BC data, the sequence of MIPS commands needed to access MIPS data projections will be presented. Then the statistical principles associated with any anomalies in the MIPS projections will be discussed. Finally an overall assessment of the MIPS application to BC data will be made.

Results

The application of MIPS to the BC data was shown to be feasible, and valid with some potential limitations of MIPS. Problems can occur with the weighted means approach when data is either missing on a significant interaction, or a significant interaction is excluded from the statistical model and the cell frequencies are nonproportional. These problems were avoided by excluding categories of data having missing data on significant interactions, and by including all significant interactions in the statistical model.

Conclusions

MIPS proved to be a valid injury projection tool when certain data parameters were met: (a) all significant effects were included in the statistical model, and (b) categories of missing data falling on significant effects were omitted. In addition, utilization of the MIPS feature for evaluating hypothetical injury projections after altering the data set demonstrated that MIPS is also a flexible injury projection tool.

3

Using the Medical Information Projection System
to Organize and Present Battle Casualty Data

INTRODUCTION

Medical resource planning for military operations requires that accurate projections be made of the anticipated incidence of battle injuries and diseases. The process of making projections consists of collecting data on previous military operations, analyzing the data, determining which predictor variables and criterion variables are informative, developing a model based on the data, and finally making projections from the model of the data.

The data evaluation process involves selecting the variables which are to be used as the criterion variables and predictor variables. For example, the criterion variables in the battle casualty (BC) data are wounded in action (WIA) and killed in action (KIA), and the predictor variables are weapon type, and ship type. Once the criterion and predictor variables have been determined, a statistical analysis will be carried out to determine if the predictor variables are meaningfully related to the criterion variables, and if the relationship between the predictors and the criterion variables varies with the different criterion variables. For example in the BC data, weapon type may be an important predictor of casualties, but the relationship between weapon type and WIA may differ from the relationship between weapon type and KIA. Results of these analyses will then provide the basis of a casualty model that can be used to make casualty estimates given various battle scenarios consisting of specific ships and weapons combinations.

The medical information projection system (MIPS) is a computerized database application program that assists the medical planner in making extrapolations from research findings [1]. This function is performed in two stages. First a medical researcher configures MIPS to access a specific data set and then imposes a statistical model on the data set. Secondly, a medical planner uses MIPS to query on specific categories in the database and receives statistically enhanced estimates for the data in the those categories. One additional function of MIPS is to allow the medical planner to add data to the current data set, and thereby carry out "what if" assessments of the data.

In order to demonstrate how MIPS achieves the functions noted above, this paper will show how MIPS was configured to access the battle casualty (BC) database [2]. The BC data used in this study reflects warships that were sunk or damaged during World War II operations. This data includes ship name, hull number, date of incident, location, weapon involved in the attack, whether the ship was sunk or damaged, number of WIA, and number of KIA.

4

The intent of this effort is fourfold: (a) to demonstrate the procedures carried out by a medical researcher to process the raw data, (b) to illustrate the procedures carried out by the researcher to configure MIPS to provide access to the statistically processed data, (c) to show the way in which a medical planner would use MIPS to access the statistically processed data, and (d) to evaluate the appropriateness of applying MIPS to the BC data.

APPLICATION OF MIPS TO BC DATA

The application of MIPS to the BC data is presented in eight sections. The first two sections provide the background information about MIPS and the battle casualty data (i.e., the overview of MIPS as a general data analytic system and the preliminary examination of the battle casualty data).

The last six sections entail a detailed sequential presentation of the steps required to make casualty projections with MIPS. These six sections consist of: (a) analyzing the BC data, (b) creating a BC data set for input in MIPS, (c) creating the MIPS statistical model for the BC data, (d) creating a MIPS parameter file from statistical model and raw data, (e) saving the MIPS parameter file, (f) using MIPS to analyze data.

Overview of MIPS As a General Data Analytic System

Three issues will be covered in this overview of MIPS. First the statistical foundation of the procedures used in MIPS will be discussed. Then a brief explanation will be given that outlines how MIPS is applied to a data set. Finally an overview of menus and options provided by MIPS will be presented.

The statistical basis for the procedures in MIPS

The statistical procedures of MIPS are based on the general linear model as it applies to multiple regression and analysis of variance[1]. Therefore, a full understanding of the following methodological description of the procedures in MIPS requires a basic knowledge of the general linear model, multiple regression (i.e., effects coding[3]), and analysis of variance (i.e., testing the significance of predictor variable effects[4]).

In analysis of variance (ANOVA) the total variance of a dependent variable is partitioned according to the factors (i.e., categorical predictor variables) and factor interactions included in the model. Each variance partition represents a unique 'effect' associated with the categorical predictors, which

5

is tested for statistical significance. The variance included in each effect is based on deviations between the cells included in the categorical predictor variables and the grand mean. By testing the variance in each effect a researcher can determine which predictor variables provide unique information about the dependent variable.

In the application of multiple regression (MR) to test the effects of categorical predictor variables on a dependent variable, the cell deviations from the grand mean for each effect are explicitly defined. For example, in the MR of a dependent variable on effects coded categorical predictor variables, statistical tests are carried out on regression weights which equal cell deviations from the grand mean. Since MR explicitly defines the cell deviations associated with each effect, the MR model can be used to separate the effects of each factor and factor interaction[1], and create estimates of cell means when nonsignificant effects are held constant.

MIPS utilizes the MR method to compute cell deviations (i.e., regression weights) between a weighted grand mean and weighted cell means, for effects which have been previously shown to be statistically significant, but MIPS does not include the computation of statistical tests on cell deviations. The exclusion of statistical tests was deemed appropriate because MIPS is designed to be used with large archival data sets that are aggregated from diverse sources, and it is expected that the parametric characteristics of archival data that are aggregated from many different sources will usually not meet the requirements of meta-analytic techniques[5] needed to estimate population variances. For example, due to data aggregation during data collection sample variances and sample sizes are often unavailable. Furthermore when the sample variances and sample sizes are provided, the reliabilities of the data collection methods for the different sources are often unknown.

In order to explain how MIPS utilizes the general linear model the following discussion is broken down into three section: utilizing the regression framework, weighted means and estimating missing data, and provisionally weighted means.

Utilizing regression framework. MIPS uses a multiple regression framework to partition data according to main effects and interaction effects. Main effects are cell deviations from the grand mean for each level of a factor. Interaction effects are cell deviations from the grand mean after the main effects of included factors are held constant. For example, if a data set of casualty counts were classified according to ship type and weapon type, the main effects would correspond to type of ship and type of weapon, and the interaction effect would correspond to the differences in casualties between ships for different types of weapons.

In MIPS, the multiple regression framework provides the basis for partitioning the data, and recombining the data when the end user queries the

6

database. When the end user queries on a particular cell of the data matrix, MIPS sums the grand mean with the effects for each factor used to define the cell. In this way MIPS delivers the cell mean for a particular cell, and/or estimates cell means for the cells either when the data is missing or some of the effects defining the cell are nonsignificant (i.e., effects not specified in the statistical model).

Weighted means and estimating missing data. MIPS uses a weighted means approach to partitioning the data. This approach can be adversely affected when the data is classified into categories of different cell frequencies. The weighted means approach gives factor levels with higher cell frequencies more influence in determining the grand mean of the criterion variable (e.g., casualty counts), and thereby influences the magnitude of individual effects (e.g., the effect of torpedoes relative to the effect of other weapons). Reconciling the influence of unequal cell frequencies on the magnitude of effect has received a lot of attention in the literature over tne years [7,8,9]. This literature has focused on the selection and development of parametric methods that enable the researcher to test the statistical significance of effects that contain unequal cell frequencies. However, the application of the weighted means method used in MIPS is only used to estimate the magnitudes of statistical effects, not to test for statistical significance. As a result, the solutions to significance testing found in the literature do not directly address the weighted means method as it is applied in MIPS.

The application of the weighted means method in MIPS becomes important when estimates of cell means are made for missing data, because MIPS uses a complex form of mean substitution for making estimates. If the cells with higher cell frequencies have higher than average casualty counts, the estimates of missing data will be higher than in the situation when the cells with higher cell frequencies have lower than average casualty counts. Giving cells with higher cell frequencies more influence over the grand mean is valid, and rational, if one assumes that the validity of the data is proportional to the number of observations in the cell (i.e., cell frequency).

Incomplete data and provisional weighted means. MIPS uses a "provisional weighted means" method for estimating cell means when the data is incomplete (i.e., data is only defined on some independent variables). The provisional weighted means method, a variant of the weighted means method, was specifically developed for use in MIPS, because archival sources often contain incomplete records of data which are missing values for some variables.

Incomplete data occurs when the data is defined according to some predictor variables, but undefined according to other predictor variables. For example, in an archival data set of casualties described according to ship type and weapon type, some records of battle casualty data may include the number of casualties and ship type, but exclude information about weapon type. If information on this ship type is incomplete, MIPS will apply the provisional

7

weighted means method when calculating cell mean **estimates.**

When incomplete data exist, the estimates of cell means based upon the weighted means approach differ from estimates of cell means based upon the provisional weighted means approach. The weighted means approach uses one mean (i.e., the grand mean) to calculated the effects of all predictor variables, while the provisional weighted means approach uses different means (i.e., provisional grand means) to estimate cell means.

The weighted means approach calculates a grand mean for all of the data, and then uses the grand mean to calculate the effects of all of the predictor variables. For example, even if the casualty data for tank landing ships does not specify the weapon type, the data from tank landing ships is used to calculate the grand mean, and the grand mean is used to calculate the relative effects of different ships and different weapons. In other words, when calculating the effects of ship type and weapon type, the casualty counts for each ship type and weapon type are compared to the grand mean.

The provisional weighted means approach will include incomplete data when calculating the grand mean, and when calculating the effect of predictor variables for which the partially classified data is defined, but not for calculating the effect of predictor variables for which the partially classified data is undefined. For example, if the casualty data for tank landing ships does not specify the weapon type, the data is defined according to ship type, but is undefined according to weapon type. Thus, the data from tank landing ships is used to calculate the grand mean for casualties, and to calculate the effect of ship type, but not to calculate the effect of weapon type. In other words, when calculating the effects of ship type the casualty counts for each ship type are compared to the grand mean, and when calculating the effects of weapon type the casualty counts for each weapon type are compared to a provisional grand mean consisting only of the data defined according to weapon type.

Applying MIPS to the BC data

Several steps are involved in applying MIPS to the BC data: analysis of raw data; creation of the MIPS raw data set; creation of the MIPS statistical model for data; merging the statistical model with the raw data, and creating a parameter file.

Overview of MIPS menus and options. MIPS has two modes of operation: researcher mode and medical planner mode. The researcher mode is a full featured implementation of MIPS that allows the researcher to configure MIPS to work with any data set, by allowing the user to define a predictive model that is used to combine input data in order to form predictions. The predictive model can consist of one or more dependent variables, and up to 100 categorical predictor variables. The medical planner mode is a restricted

8

implementation of MIPS that enables the medical planner to query a database and
add additional raw data to the database, but does not permit the user to modify
the model used to form predictions.

1. The medical planner mode: The medical planner mode provides access to
two menus of options. The first menu is the primary option menu, and the
second menu is the query database menu.

The primary option menu shown in Figure 1 provides the user with three
procedure options, and one system option that allows the user to 'Quit MIPS'.
Before the user can gain access to the second or third procedure options they
must select the first option and specify a parameter file. Once a parameter
file has been specified, the "query database" option, and the "add raw data"
options become available.

The query database option leads to a second menu, while the add raw data
invokes a closed loop sequence of prompts that allow the user to add raw data.
Adding raw data permits the end user to change the estimated cell means (i.e.,
the parameter estimates), and thereby evaluate the effect of hypothetical data
sets.

```
----------------------------------------------------------
        Medical Information Projection System

        Read parameter file    1
        Query Database         n/a
        Add Raw Data           n/a
        Quit MIPS              CR
----------------------------------------------------------
```

Figure 1. The MIPS primary option menu

Selecting the "Query Database" option from the primary option menu results
in the "Query Database Menu" (soe Figure 2). This menu provides two ways to
query the data base: cross-tabulate query and casewise query. This menu also
provides a procedure for creating casewise query files that are used by the
casewise query procedure to define the cases being queried.

The cross-tabulate query option allows the user to specify the categories
of interest for each predictor variable in the data set. Then the program
builds a cross classification table for the combination of the levels of each
variable specified by the user. For example, if the user chooses battleships
and cruisers from the ship type variable, and torpedoes and bombs from the
weapon type variable, the program will give a casualty estimate for battleships
and cruisers hit by torpedoes, and battleships and cruisers hit by bombs.

9

The casewise query option allows the user to specify single cases defined by more than one variable. For example, the user could request two battle cases consisting of a destroyer hit by a torpedo, and a cruiser hit by a bomb, and exclude information about destroyers hit by bombs, and cruisers hit by torpedoes. In other words, MIPS does not create a cross classification of the variables used in the query. This option allows the user to acquire casualty estimates for specific convoy scenarios in which different ships are struck by different weapons.

Option three gives the user the ability to create "casewise" query files. These files are created before the "casewise query option" is invoked, because these files contain the information that defines the cases on which the query is performed. As a result, a number of casewise query files can be defined, saved, and then accessed by the "casewise query option". For example, several casewise query files could be created for the BC data that define different battle scenarios. Then casualty estimates for different battle scenarios can be requested by invoking the "casewise query option" and specifying the casewise query file corresponding to the battle scenarios of interest.

```
-------------------------------------------------
      Query DataBase

          Cross Tabulate Query          1
          CaseWise Query                2
          Build CaseWise Query file     3
          Quit Query                    CR
-------------------------------------------------
```

Figure 2. The MIPS define and maintain statistical model

2. The researcher mode: The researcher mode gives the user expanded access to the MIPS procedures. This mode provides access to four menus of options: (a) the primary option menu, (b) the build parameter file menu, (c) the define and maintain statistical model menu, and (d) the query database menu.

The primary option menu shown in Figure 3 initially provides the researcher with two options that allow MIPS to interface with a database. If a data base and statistical model have already been used to configure a parameter file, the researcher can access the database by selecting option one. If no parameter files exist, the researcher can build a new parameter file by selecting option two. The process of building a parameter file is described in detail in subsequent sections of this document.

10

Once a parameter file has been selected by accessing the first option or by building a new parameter file, the "query database" option, and the "add raw data" option will be made available.

```
-------------------------------------------------------
            Medical Information Projection System

            Read parameter file    1
            Build parameter file   2
            Query Database         n/a
            Add Raw Data           n/a
            Quit MIPS              CR
-------------------------------------------------------
```

Figure 3. The MIPS primary option menu in researcher mode

If the researcher initially selects the "build parameter file option", MIPS provides an alternate menu (see Figure 4). Although each option will be explained in greater detail in a following section of this paper, a brief description will be given at this point.

The process of building a parameter model consists of four steps. First, the variables need to be defined, by selecting options one, two, or three. Second, the statistical model needs to be defined by selecting option four. The data is then read and the MIPS program calculates estimates of cell means for each variable, and variable interaction, defined in the statistical model (i.e., parameter estimates for statistical effects). Finally, the parameter estimates are written to a disk file containing the MIPS database.

```
-------------------------------------------------------
        Build Parameter Model

        1    Specify variable definitions
        2    Read variable definitions
        3    Read var. def. & statistical model
        4    Define and maintain statistical model
        5    Read data and create parameter file
        6    Write parameter file
        CR   Quit build
-------------------------------------------------------
```

Figure 4. The MIPS build parameter model menu

When the "Define and maintain statistical model" option is selected from

11

the "build parameter file" menu, MIPS presents the "Define and maintain statistical model" menu (see Figure 5). This menu contains two options that enable the definition of a statistical model. Options one and three are for information only, and give the researcher the opportunity to review the variable definitions, and current statistical model. Option two provides the capability to define each statistical effect one at a time. Options four and five allow the researcher to delete effects, and to process the deletions without having to return to the previous menu (i.e., "Build Parameter Model" menu). Option 6 provides the capability to read the statistical model from the summary page of a statistical analysis using the SAS GLM procedure[3].

Define and Maintain Statistical Model

1	List classifying variables and number of categories
2	Define statistical effects
3	List effects
4	Delete an effect
5	Sort effects and process deletions
6	Get SAS Model
CR	Quit define and maintain statistical model

Figure 5. The MIPS define and maintain statistical model

Preliminary Examination of the Battle Casualty Data

A preliminary examination of the battle casualty (BC) data is needed to establish a data framework that can be used in subsequent discussions of the application of MIPS to the BC data. The BC data in the current study can be broken down according to two predictor variables and one criterion classification variable. Note that criterion classification variables occur in other statistical procedures such as in a repeated measures analysis of variance (e.g., time), or in an inter-rater reliability analysis of variance (e.g., rater). The two predictor variables consist of: 'type of ship', and 'type of weapon'. The criterion classification variable is 'type of casualty'. The criterion classification variable is included because there are two measures of casualty counts (i.e., WIA, and KIA), and it is desirable to compare the levels of WIA and KIA.

These two predictor variables and one criterion classification variable form a set of three independent variables in the statistical model. These three independent variables can be evaluated according to each variable

12

individually, and in combination by cross classifying the data according to various combinations of the variables. Consequently, these three variables can be combined into seven orthogonal statistical effects. See Table 1 below for a listing of all of the possible statistical effects in the BC data.

Table 1. Ways of Classifying the WWII Battle Casualty Data

    1. Ship-type (collapsed across casualty-type and weapon)
    2. Weapon (collapsed across casualty-type and ship-type)
    3. Casualty-type (collapsed across ship-type and weapon)
    4. Ship-type X Weapon (collapsed across casualty-type)
    5. Ship-type X Casualty-type (collapsed across weapon)
    6. Weapon X Casualty-type (collapsed across ship-type)
    7. Ship-type X Weapon X Casualty-Type

Each way of classifying the BC data (i.e., statistical effect) provides unique information about the casualty counts. However, the information provided by a particular classification may not be useful if differences in the cell means of the classification are negligible. For example, if the difference in casualty counts between ship types is negligible, the ship type distinction is not useful in predicting casualty counts. If the ship-type classification reveals meaningfully large and reliable differences in casualty counts, the ship-type variable will be useful in predicting casualty counts. The reliability of the casualty count differences will be evaluated statistically by comparing the casualty count differences between ship-types with the variability of casualty counts within ship-types. If the casualty rate differences between ship-types is relatively large compared to the casualty rate differences within ship-types, the ship-type effect should be reliable.

It is important to note that the seven ways of classifying this data set fall into two general categories. The first category consists of "main" effects for each of the three variables, and includes the first three ways of classifying the data shown in Table 1. A main effect of a variable represents the differences between the category means included in the variable. An interaction effect of a combination of variables represents the differences between the category means included in a cross classification of the variables involved in the interaction. For example, the Ship-type X Weapon classification (i.e., interaction) represents the difference between the means included in the cross classification. This interaction reveals the extent to which the casualty counts for different weapons vary according to ship-type. If the weapon casualty counts in this cross classification are found to differ significantly between ships, this information will be useful in predicting casualty counts.

13

Analysis of BC Data

The analysis of the data consists of two steps. First, a set of meaningful variables needs to be selected. Then a statistical analysis provides a basis for choosing variables (i.e., main effects) and variable interactions to be included in the MIPS model.

Selecting a set of meaningful variables

Meaningful independent variables will usually be those variables which represent significant main effects. However, independent variables which are not statistically significant may still be meaningful if the variables have been adopted over time because they are useful in describing the data. For example, the main effect for casualty type was not statistically significant for the present shipboard BC data, but the distinction of WIA and KIA is very important to medical planners. Therefore, the user is given access to individual WIA and KIA estimates, even though the casualty type main effect is not included in the statistical model.

Determining significant effects and interactions

The set of meaningful variables will be evaluated in a statistical analysis, and the significant main effects and interaction effects resulting from this analysis will be included in the statistical model used by MIPS. Nonsignificant effects are excluded from the statistical model because they represent nonsignificant differences between the cell means of the main effect or interaction effect, and nonsignificant differences between the cell means are probably random differences (i.e., noise) which should not be allowed to influence cell mean estimates.

When nonsignificant effects are omitted the variables defining the effect can be used to classify the data, but the cell mean estimates for the nonsignificant variable will be equal. For example, if the main effect of casualty type is omitted from the statistical model the individual estimates of WIA and KIA (i.e., collapsed across ship and weapon type) will be equal. As a result, the WIA and KIA estimates will not reflect the potentially random differences occurring in the data.

Omitting a nonsignificant interaction term has a similar effect on the query process and the cell mean estimation method as the omission of a main effect. The user can still query the database at the level of the interaction, but the cell mean estimates will not reflect the random variation associated with the omitted interaction effect. For example, if the ship by weapon interaction was not significant, and the interaction effect was omitted from the statistical model, the user could still request cell mean estimates for the

ship by weapon cross classification. In this situation the cell mean estimates would only reflect the addition of the ship effect and the weapon effect; differences in the cell means of the raw data for the ship by weapon cross classification beyond the additive contribution of each variable are not included in cell mean estimates.

The analysis of the BC data set resulted in four significant effects (i.e., four factors and/or factor interactions). See Table 2 below. Several effects were not statistically significant. The casualty type effect, for example, was not significant because on average the number of WIA and KIA are the same for all of the WWII shipboard battle incidents. Consequently, the MIPS model should not include this effect because it will introduce noise in the casualty estimates.

Table 2. Significant effects in the WWII Battle Casualty Data

1. Ship-type (collapsed across casualty-type and weapon)
2. Weapon (collapsed across casualty-type and ship-type)
3. Ship-type X Weapon (collapsed across casualty-type)
4. Weapon X Casualty-type (collapsed across ship-type)

Creation of BC Data Set For Input In MIPS

The creation of a data set consists of determining what variables need to be included, determining how the observations will be weighted, and assembling the data into an ASCII data file.

Determination of classification variables

Determining which classification variables will be included in the data set follows from the statistical analysis. All variables that were included in significant effects will be used. Other variables can be included if they have historical significance, but they will only be used to query the database and not in the statistical model. The BC data set used with MIPS consists of the variables that were either significant main effects or part of significant interactions: (a) ship type, (b) weapon type, and (c) casualty type.

Determination of observation weight

MIPS, like many regression programs allows the user to weight the observations in the raw data set. In the BC data each incident was equally weighted with a weight of one. A data set could use case weights greater than one if the raw data were aggregated. For example, if the BC data were

15

aggregated according to ship type and weapon type, one mean would be available for each ship and weapon combination, and that mean could be given a weight reflective of the number of observations used to calculate the mean.

Data set format

The raw data set read by MIPS consists of two header lines followed by the raw data. The first header line specifies the number of independent variables, which includes predictor variables (i.e., ship type and weapon type) and criterion classification variables (i.e., casualty type). A criterion classification variable is required when multiple criterion variables are included in the database. The number of independent variables is located in columns 1-4. The second header line specifies the FORTRAN format of the independent variables, the criterion data variable, and the weight for each case. The criterion data variable which contains the raw data (i.e., casualty counts) is different from the criterion classification variable which contains an integer code for the individual criterion variables (i.e., WIA or KIA). The data format is located in columns 1-80.

The raw data that follows the two header lines consists of one line of data for each criterion variable included in the database. Therefore, the BC data includes two lines of data for each battle incident. One line for WIA, and one line for KIA. For each line of the BC data the first number is ship type, the second number is weapon type, the third number is casualty type, the fourth number is the casualty count, and the fifth number is the data weight. In the BC database the weight of one indicates that the casualty count is based on one battle incident. If the data were aggregated the casualty count would equal the sum for the aggregation, and the weight would equal the number of incidents for the aggregation. See Figure 6 showing both header lines and the first two lines of data. The hull number and date of incident are also included on each case of data for reference purposes.

```
------------------------------------
  4
(I3,I2,I2,2F4.0)
  1 1 1  20   1  AM-123   6 10 44
  1 1 2   3   1  AM-123   6 10 44
------------------------------------
```

Figure 6. Two data file header lines with first two lines of data. The first line specifies the number of independent variables. The second line specifies the FORTRAN format for the following data. The next two lines of data contain one line of WIA data and one line of KIA data.

Creating a MIPS Database

The creation of a MIPS database can only be accomplished when MIPS has been invoked in research mode. The research mode is made available to the MIPS user when the default directory contains a file called 'research.tst'.

Once MIPS is invoked in research mode the process of creating a MIPS database involves accessing several MIPS options in order to configure MIPS for the data being analyzed. First, the researcher must determine how the independent variables (i.e., predictor, and criterion classification variables) will be labeled, and then use the MIPS option "Specify variable definitions" to configure MIPS. Next, the researcher determines which statistical effects will be included, and then uses the MIPS option "Define statistical effects" to configure MIPS. Next, the raw data will be read in to create a MIPS parameter file by using the MIPS option "Read data and create parameter file". Finally, the researcher will save the new database by using the MIPS option "Write parameter file". A detailed explanation of these steps follows.

Determination and Specification of Variable Definitions

The determination and specification of variable definitions consists of establishing labels for the variables included in the MIPS raw data set, and then utilizing a MIPS procedure to specify the variable definitions.

Determination of variable labels. The determination of the variable labels is achieved in four steps: (a) establishing the data to be used with MIPS, and the variables used to describe the data. (b) the determination of the number of variable levels, (c) the determination of logical limits for data, and (d) the determination of criterion classification variables.

1. The data and variables to be incorporated in the MIPS database was determined in the analysis phase. The BC data set was classified according to (a) ship type, (b) weapon type, and (c) casualty type. For the purpose of encoding these variables for use in MIPS the variable names were shortened to eight characters. The new variable names are shiptype, weapon, and castype.

2. The Determination of the number of variable levels for each classification variable and the label for each level is based on the categories included in each variable. For example, there are four categories of weapon type: BOMB, GUNFIRE, KAMIKAZE, and TORPEDO. The category names are used as level labels.

3. Determination of logical limits for data. Since MIPS uses a regression model to project estimates of the data, some estimates may exceed logical limits of the variables that define the data set. Logical limits may exist for the levels of a particular variable. For example, the number of

17

casualties for a ship can not exceed the number of personnel aboard. In the
BC data logical limits were established for the levels of the 'SHIPTYPE'. For
example, the logical limits for destroyers and cruisers equal 382, and 1618
respectively.

4.   Determination of criterion classification variables.   The term
criterion classification variable is used here to describe variables such as
casualty type which encode multiple criterion variables that are available for
each data observation (e.g., each battle incident). In other words, casualty
type is a criterion classification variable because both WIA and KIA counts are
available for each battle incident included in the BC data set. With the
incorporation of multiple observations on individual data observations in one
data set, the analysis methods used by MIPS become analogous to certain
multivariate statistics. For example, the method for criterion classification
variables used in MIPS is comparable to the inter-rater reliability analysis
of variance.   Tests of inter-rater reliability assess the degree to which
judges agree in their ratings of the same subjects. In the case of the BC data
the agreement between WIA and KIA as indicators of casualty levels, is
evaluated for the same battle incidents.

Because casualty type is a criterion classification variable, MIPS forces
the user to include at least one casualty type (i.e., WIA or KIA) in all
database queries.

   Specification of variable labels.   Variables labels are specified by
selecting the MIPS option called 'Specify variable definitions' (Figure 7).
In order to gain access to this option, the 'build parameter file' option must
be selected from the primary option menu (Figure 3).

```
------------------------------------------------------------

     Build Parameter Model

          1    Specify variable definitions
          2    Read variable definitions
          3    Read var. def. & statistical model
          4    Define and maintain statistical model
          5    Read data and create parameter file
          6    Write parameter file
          CR   Quit build

------------------------------------------------------------
```

Figure 7.   Build parameter file menu.

Determining and specifying statistical model in MIPS

The determination and specification of the statistical model in MIPS consists of: (a) analyzing the data to determine which statistical effects are significant, and (b) utilizing a MIPS procedure to specify the statistical model.

Determining the statistical model to be used in MIPS. Once the variables of the data set have been defined MIPS, allows the user to define the statistical model. The statistical model consists of all effects that were found to account for a significant amount of the variance in the dependent variable. In the BC data the significant effects are as follows: SHIPTYPE, WEAPON, WEAPON BY SHIPTYPE, AND WEAPON BY CASTYPE.

Specifying the statistical model in MIPS. The 'Define statistical effects' option is accessed from the 'Define and Maintain Statistical Model' menu (Figure 8), which is an option of the 'Build parameter file' menu.

```
------------------------------------------------------------------
Define and Maintain Statistical Model

    1    List classifying variables and number of categories
    2    Define statistical effects
    3    List effects
    4    Delete an effect
    5    Sort effects and process deletions
    6    Get SAS Model
    CR   Quit create and maintain model and, sort and delete.
------------------------------------------------------------------
```

Figure 8. Define and Maintain Statistical Model

Creating Parameter File From Statistical Model and Raw Data

Once the statistical model has been specified, MIPS will allow the researcher to read in the raw data and create a MIPS parameter file. The researcher needs to select the 'Read data and create parameter file' option from the 'Build Parameter Model' menu.

One of the prompts during the 'Read data and create parameter file' option is to input the label for the cell mean estimates generated by MIPS. The appropriate label must describe the criterion variables and unit of measurement. The criterion variables for the BC data are measures of

19

casualties. The unit of measurement are casualty counts per battle incident, as opposed to casualty percentages (i.e., percent of the ships complement injured per incident). Given the name of the criterion variable and the unit of measurement the casualty estimates made by MIPS are labeled as "casualty counts".

Saving MIPS Parameter File

Once MIPS reads the raw data and calculates parameter estimates, the researcher can save the parameter file. This is achieved by selecting the 'Write parameter file' option from the 'Build Parameter Model' menu.

Utilizing MIPS to Analyze Data

Utilizing MIPS to analyze data consists of four basic activities: (a) reading a parameter file, (b) querying a MIPS database, (c) executing "what-if" scenarios, and (d) quitting MIPS and printing a listing file.

Reading a parameter file

Once a MIPS parameter file has been created, MIPS can read the file to allow the user to query the data. Reading a parameter file is done by selecting the 'Read parameter file' option from the 'Primary option menu'.

Querying a MIPS database

Once a MIPS parameter file has been read, the user may query the data by selecting the 'Cross-tabulate query' option, or the 'Casewise query' option. These MIPS options are accessed through the 'Query Database' Menu (see Figure 2).

Using the 'Cross-tabulate query' option. When the 'Cross Tabulate Query' option has been selected, the user can specify levels of each factor on which the user needs information. MIPS then generates a cross-tabulated list of all variable levels requested. For example, in the BC data if the user requested two ships (i.e., cruisers, and carriers), and two weapons (i.e., bombs, torpedoes) and the WIA casualty type MIPS would generate a list consisting of WIA casualty count projections for cruisers hit with either bombs or torpedoes, and projections for carriers hit with either bombs or torpedoes. See Figure 9 for an example of the MIPS listing of casualty count estimates.

20

```
    ---------------------------------------------------------
    Category Selected:
    SHIP      WEAPON    CASTYPE   Cell Mean
    ---------------------------------------------------------
    CA        BOMB      WIA          14.168
    CV        BOMB      WIA          73.177
    CA        TORPEDO   WIA          82.512
    CV        TORPEDO   WIA          22.779
    ---------------------------------------------------------
    Total Casualty Count = 192.637
    Average Casualty Count =  48.159
    Number of Casualty Count projections =    4
    ---------------------------------------------------------
```

Figure 9. MIPS cross-classification listing of WIA casualty estimates for two ships (i.e., CA, and CL), and two weapons (i.e., torpedo, and bomb).

Using 'Casewise query' option. When the 'Casewise query' option has been selected MIPS reads a 'casewise query file' that contains one or more cases of interest. For example, a casewise query file could be created to represent WIAs for a convoy group of one cruiser hit with bombs, one carrier hit with torpedoes, and one destroyer hit with torpedoes. Note that the list of cell mean estimates does not represent a cross classification of bombs and torpedoes with cruisers, carriers, and destroyers. See Figure 10 for an example of the MIPS output for this convoy example.

Figure 10. MIPS WIA estimates for a casewise query file representing a convoy of one cruiser hit with bombs, one carrier hit with torpedoes, and one destroyer hit with torpedoes.

```
    ---------------------------------------------------------
    Category Selected:
    SHIP      WEAPON    CASTYPE   Cell Mean
    ---------------------------------------------------------
    CA        BOMB      WIA          14.168
    CV        TORPEDO   WIA          22.779
    DD        TORPEDO   WIA          23.144
    ---------------------------------------------------------
    Total Casualty Count =  60.092
    Average Casualty Count =  20.031
    Number of Casualty Count projections =    3
    ---------------------------------------------------------
```

21

Executing a "what-if" scenario

Executing a "what-if" scenario requires that the user identify the effect that is to be manipulated. For example, the user may want to see the overall effect of increasing the impact of torpedoes. In order to modify the torpedo effect, select the 'Add Raw Data' option from the primary option menu (Figure 1). Selecting this option will initiate a closed loop series of prompts: (a) requesting the classification of the data on the variables defining the data, (b) the numerator for the cell mean calculation, and (c) the denominator (i.e., weight) for the cell mean calculation. After the new information is verified, MIPS adds the new data to the existing model and updates the parameter estimates. After the new data is incorporated, the user can query the database to see the effects of the new information.

For example, the effect of torpedoes could be modified by adding cases where the torpedo effect was doubled. In order to determine the cell value and weight of the case to be added, one needs to determine the current torpedo effect, and the number of observations included in this effect (i.e., weight). By evaluating the MIPS parameter file (Appendix A), the user can determine that the number of torpedo casualties for 210 battle incidents equals 10,433. This translates to 49.7 casualties per incident. Given this information, one can determine the casualty count and incident count (i.e., weight) that is needed to be added in order to simulate some hypothetical casualty rate.

If one wanted to see the effect of an additional 210 cases with twice the casualty rate, the user would add a case of data having 20,866 casualties with a weight of 210. See Figure 11 for the WIA estimates for the convoy in Figure 10, after the new data has been added.

```
-----------------------------------------------------------
     Category Selected:
     SHIP       WEAPON     CASTYPE  Cell Mean
     --------------------------------------------------
     CA         BOMB       WIA         14.168
     CV         TORPEDO    WIA         46.667
     DD         TORPEDO    WIA         47.031
     --------------------------------------------------
     Total Casualty Count = 107.866
     Average Casualty Count =  35.955
     Number of Casualty Count projections =   3
-----------------------------------------------------------
```

Figure 11. MIPS WIA estimates for a casewise query file after the effect of torpedoes was modified by adding 210 cases with a mean of 97 total casualties per battle incident.

22

It should be noted that future implementations of MIPS will be designed to allow the user to directly apply arithmetic operations on each level of each effect without being concerned with weight of the current estimate. For example, the user could multiply the torpedo effect by 2.0, or increment the bomb effect by 100.

Quitting MIPS and printing listing file

Exiting from MIPS is achieved by answering all MIPS menu prompts with a carriage return. In this way the user can back out to the primary option menu and quit MIPS.

ANALYSIS OF PROJECTED MEANS USING MIPS

Assessment of Assumptions

As with any data analytic system certain assumptions must be met or inference problems can occur. With MIPS, inference problems can occur if any of three assumptions are not met.

The first assumption states that the categories included in significant effects must not have empty cells.

The second assumption states that statistical effects that are omitted from the statistical model must not have nonproportional cell frequencies.

The third assumption states that the cross classification of variables must not include illogical categories.

Empty cells on significant effects

Estimating cell means for empty cells is possible if the effects defining the data in the empty cell are not significant. Ward & Jennings[10] demonstrate how the side conditions specified by ANOVA permit the estimation of the cell mean for an empty cell in the cross classification of two factors. But this estimation method is based on the assumption that the interaction defining the empty cell is nonsignificant.

When a factor is significant, empty cells on that factor probably have cell means that are significantly different from the grand mean. Consequently, when the categories of a predictor variable such as ship type show that different ships have significantly different casualty counts, it is not

23

statistically advisable to use the grand mean as an estimate for the casualty frequencies for ships not included in the data set. Casualty estimates for ships on which data is missing can only be made if these ships can be associated with other ship types for which data is available. The process of comparing ships having data with ships not having data must be done by an informed person on a case by case basis. This type of inference is outside the design parameters established for the current version of MIPS.

The issue of empty cells on significant effects becomes more complex when empty cells occur on significant interaction effects. Interactions between variables are analogous to the cross classification of variables. For example, the interaction between ship type and weapon type reveals a weapon profile of casualties for each individual ship. A significant interaction would indicate that the weapon profile of casualties differs between ships. Missing data on this interaction could occur when data was missing for some ship by weapon combinations. Therefore, though data is available for all ship types and all weapons, if the interaction is significant all ship and weapon combinations having empty cells must be omitted from the database. However, if the interaction is not significant, MIPS can be used to generate cell mean estimates for empty cells on the cross classification.

Omitting effects that contain unequal cell frequencies

Cell frequencies are a measure of the number of observations falling in the categories included within a factor or factor interaction. Inference errors can potentially occur when effects are omitted and the categories included in the omitted effect have unequal cell frequencies, because although the effect is excluded the unequal cell frequencies will cause the different categories of the effect to have a differential impact on estimates of cell means. This is problematic because the effect was omitted as it would have contributed noise in estimates of cell means, but when the effect is excluded it still has an effect on cell mean estimates due to unequal cell frequencies.

The most simple example of the omitted effect containing unequal cell frequencies occurs when a main effect is nonsignificant, but the categories of the main effect have unequal cell frequencies. For example, in the BC data there are different cell frequencies for each weapon in the weapon-type factor. If weapon-type was not a significant predictor of casualties, it would have been excluded from the statistical model in order to reduce unwanted noise in casualty predictions. However, since weapon-type has unequal cell frequency the different weapon types would impact cell mean estimates. In the current data set weapon-type and ship-type had unequal cell frequencies, but since they were significant casualty predictors they were included in the statistical model, consequently no problem occurred. Casualty type on the other hand, was not a statistically significant classifier of casualties, but since there are an equal number of WIA and KIA observations problems associated with unequal

24

cell frequencies do not apply.

It should be noted that, in some situations, the effects due to unequal cell frequencies are meaningful, not undesirable. Unequal cell frequencies are meaningful if the number of observations within cells are associated with the topic under study. For example, if the number of casualties occurring per battle incident were related to the number of recorded incidents, cell frequencies could provide useful information in the prediction of casualties. This is not the case with the current data under study.

The problem of unequal cell frequencies is more complicated for interaction effects than main effects. An interaction with unequal cell frequencies may be safely omitted if the interaction cell frequencies are proportional to the cell frequencies of the included lower order effects, and those lower order effects are included in the statistical model. For example, if the cross classification of ship and weapon type had unequal cell frequencies, these cell frequencies will be either proportional or nonproportional to the cell frequencies of ship type and weapon type. If all ship types had the same cell frequencies, and only torpedoes had lower cell frequencies than other weapons, then the torpedo cell frequencies would be low for all ship types and the cell frequencies of the weapon by ship type cross classification would be considered proportional to the cell frequencies of weapon type and ship type. Nonproportional cell frequencies would occur if the cell frequencies for the various weapons differed among ship types. To reiterate, an estimation problem may occur with MIPS if an interaction term is omitted, and cell frequencies for the interaction term are not proportional to the cell frequencies of the included lower order effects.

Three examples in Table 3, 4, and 5 illustrate the problem of omitting interaction terms when the data meet the following two conditions.

1. The interaction term for the cell means is not significant (i.e., cell means are proportional to the means of included lower order effects),

2. The cell frequencies effect is significant (i.e., the cell frequencies are nonproportional to the cell frequencies of the included lower order effects).

In each example the interaction term for the cell means is not significant. However, the cell frequency condition varies from one example to the next example. The interaction effect is represented by the cross classification of the two ship types with the two weapon types. The main effect for weapon types and ship type is represented by the marginal means. The main effect for weapon type is found in the row for all ships, and the main effect for ship type is found in the column for all weapons. The weighted means estimate for the interaction cells is equal to the sum of the grand mean plus the appropriate ship effect (i.e., the ship mean minus the grand mean)

plus the appropriate weapon effect (i.e., the weapon mean minus the grand mean).

Each example provides: (a) the cell mean, (b) the cell frequency, and (c) the weighted means estimate of the cell mean when the interaction term is omitted from the statistical model. The example in Table 3 shows the weighted means estimate of the cell means when cell frequencies are equivalent; therefore the cell frequency effect is not important. Table 4 shows the weighted means estimate of the cell means when the cell frequencies are unequal but proportional; therefore, the cell frequency effect is not important. Table 5 shows the weighted means estimate of the cell mean when the cell frequencies are unequal and nonproportional, therefore, the cell frequency effect introduces errors to the estimation of the cell means.

The lack of an interaction effect for the means is evident because the differences between interaction means equals the difference between main effect means; therefore, the difference between interaction means can be explained by the differences between main effect means. For example, the difference between the two ship types for 'weapon 1' equals the difference between the two ship types for 'weapon 2', and 'All weapons'. This indicates a nonsignificant interaction and that the difference between the two ships in the cross classification can be explained by the main effect of ship type. The same logic can be used to illustrate the cell frequency interactions.

26

Table 3. A hypothetical two factor cross classification of weapon-type and ship-type casualty counts with: (a) equal cell frequencies, (b) significant main effects, and (c) a nonsignificant interaction. Predicted values are based on a weighted means main effect regression model. In this example the cell frequencies for the interaction term are proportional to the cell frequencies of the main effects. As a result, the weighted means estimates of the cell means equal the observed cell means.

Weapon-type

	weapon 1	weapon 2	All Weapons
Ship-type			
ship 1	mean= 5 frequency= 1 estimate= 5	mean= 25 frequency= 1 estimate= 25	mean= 15 frequency= 2
ship 2	mean= 25 frequency= 1 estimate= 25	mean= 45 frequency= 1 estimate= 45	mean= 35 frequency= 2
All Ships	mean= 15 frequency= 2	mean= 35 frequency= 2	mean= 25 frequency= 4

27

Table 4. A hypothetical two factor cross classification of weapon-type and ship-type casualty counts with: (a) proportional cell frequencies, (b) significant main effects, and (c) nonsignificant interactions. Predicted values are based on a weighted means main effect regression model. In this example the interaction cell frequencies are proportional to the main effect cell frequencies. Consequently, the weighted means estimates of the cell means equal the observed cell means.

Weapon-type

	weapon 1	weapon 2	All Weapons
Ship-type			
ship 1	mean= 5 frequency= 1 estimate= 5	mean= 25 frequency= 10 estimate= 25	mean= 23.2 frequency= 11
ship 2	mean= 25 frequency=100 estimate= 25	mean= 45 frequency=1000 estimate= 45	mean= 43.2 frequency=1100
All Ships	mean= 24.8 frequency=101	mean= 44.8 frequency=1010	mean= 43.0 frequency=1111

Table 5. A hypothetical two factor cross classification of weapon-type and ship-type casualty counts with: (a) nonproportional cell frequencies, (b) significant main effects, and (c) nonsignificant interactions. Predicted values are based on a weighted means main effect regression model. In this example the interaction cell frequencies are nonproportional to the main effect cell frequencies, therefore, the weighted means estimates of the cell means do not equal the observed cell means. It should be noted that the nonproportional cell frequencies caused the main effect means to be nonproportional to the interaction means.

Weapon-type

Ship-type	weapon 1	weapon 2	All Weapons
ship 1	mean= 5 frequency= 1 estimate= -3.3	mean= 25 frequency= 1 estimate= 16.8	mean= 15 frequency= 2
ship 2	mean= 25 frequency=100 estimate= 28.9	mean= 45 frequency=1000 estimate= 45	mean= 43.2 frequency=1100
All Ships	mean= 28.8 frequency=101	mean= 45 frequency=1001	mean= 43.1 frequency=1102

Illogical categories for the cross classification of variables

Illogical categories occur when the cross classification of variables yield categories for which data is not possible. Illogical categories can occur for some variables used to define battle casualties that were not included in the data used in this paper. For example, two problematic BC variables are injury type (e.g., burns, and sprains), and injury location (e.g., leg, and head). It is readily evident that the cross classification of injury type and injury location results in some illogical categories such as a head sprain.

When MIPS works with data sets that contain illogical categories MIPS deals with the category as if the data could exist, but the data is simply missing. As a result, the program will estimate a value for the missing data. In the example above MIPS would yield a casualty count estimate for head sprains.

29

Assessment of Assumptions in Regard to the BC Data

The application of MIPS to the BC data needs to be evaluated in regard to the assumptions: (a) data must be available for all categories included in significant effects, (b) omitted effects must not contain nonproportional cell frequencies, and (c) the cross classification of predictor variables must not result in illogical categories.

Empty cells on significant effects

The BC data had missing data in the ship-type by weapon-type cross classification, and the interaction between ship and weapon was statistically significant. Because of the limitation of not having empty cells on significant effects, several ship types (i.e., PT, YMS, and DMS) and two weapon types (i.e., mines, and multiple), were omitted from the data set.

Omitting effects that contain unequal cell frequencies

The MIPS model of the BC data did not exclude any variables or cross classification of variables that had unequal and nonproportional cell frequencies. The only interactions that were omitted from the model were casualty type interactions, and the frequencies of WIA and KIA were always equivalent. Therefore, the cell frequencies for interactions with casualty type are proportional to the cell frequencies of the effects interacting with casualty type, and all of these effects were included in the statistical model.

Illogical categories for the cross classification of variables

All of the cross classifications possible for ship type, weapon type, and casualty type represented plausible incidents. Therefore, illogical categories do not occur in the BC data.

CONCLUSION

MIPS proved to be a valid injury projection tool when certain statistical underpinnings were met.

The main benefits of using MIPS with the BC data derive from the ability to quickly determine the casualties of specific convoy groups being hit by specific weapon types. Present day projections may need to be modified according to changes in ships and weapon systems. However, the data delivered by MIPS may be used as a basis for current day extrapolations, or for a quick check of casualty estimates derived by other means.

30

In addition, MIPS allows the user to modify the database in order to evaluate "what if" scenarios. This powerful feature allows planners to add real or hypothetical battle incidents to the data base and see how they effect previous estimates of casualty counts.

REFERENCES

1. **Pugh, William:** <u>Use of Artificial Intelligence Techniques for Medical Readiness Planning</u>. Technical Report No. 88-33. Naval Health Research Center, 1988.

2. **Blood, Christopher:** <u>Analyses of Battle Casualties by Weapon Type Aboard U.S. Navy Warships</u>. Technical Report No. 90-1. Naval Health Research Center, 1991.

3. **Cohen, J., & Cohen, P.:** <u>Applied Multiple Regression for the Behavioral Sciences</u>. New York: Wiley, 1975.

4. **Winer, B.J.:** <u>Statistical Principles in Experimental Design</u>. 2nd Ed. New York: McGraw-Hill, 1971.

5. **Hunter, J. E., Schmidt, F.L., & Jackson, G.G.:** <u>Meta-analysis: Cumulating research findings across studies</u>. Beverly Hills, Ca: Sage, 1982.

6. **SAS Institute:** <u>SAS User's Guide: Statistics, Version 5 Edition</u>. Cary, NC: SAS Institute Inc, 1985.

7. **Carlson, J. E., Timm, N. H.:** Analysis of nonorthogonal fixed-effects designs. <u>Psychological Bulletin</u>, 1974, 81, 9, pps. 563-570.

8. **Overall, J. E., Spiegel, D. K.:** Concerning least squares analysis of experimental data. <u>Psychological Bulletin</u>, 1969, 72, 5, pps. 311-322.

9. **Overall, J. E., Spiegel, D. K., Cohen, J.:** Equivalence of orthogonal and nonorthogonal analysis of variance. <u>Psychological Bulletin</u>, 1975, 82, 2, pps. 182-186.

10. **Ward, J. H., Jennings, E.:** <u>Introduction to linear Models</u>. New Jersey: Prentice-Hall. pps. 231-237.

Appendix A: MIPS Battle Casualty parameter file

The Mips parameter file consists of four arrays of data. See Figure A1 below.

The first array includes information on each variable used to describe the data set. The information consists of the variable number, the number of levels included in the variable, the name of the variable, and a code to indicate if the variable is a predictor variable or a criterion classification variable.

The second array includes information on the levels of each variable used to describe the data set. The information consists of the variable number, the level number, the variable level label, and a logical upper limit for data falling in that category.

The third array includes information on the statistical model used to describe the data. This array includes three lines for each statistical effect defined in the statistical model. One line is used to specify each variable included in the effect. Three lines are used to define each effect because the current version of MIPS allows for a maximum of three variables to be included in each effect. The information contained on each line describing the effect consists of the number of the statistical effect, the line number used to describe each included variable, the variable number of the variable included in the effect, an index number used by MIPS to locate category cells included in the effect, and the name of the variable included in the effect.

The fourth array lists data on each category included in each statistical effect included in the statistical model. The information consists of the effect number, the effect category number, the numerator for the data in that category (i.e., the sum of the casualties for the category), the denominator for the data in that category (i.e., the frequency of observations for the category), a weighted means estimate of the effect for that category, a provisionally weighted means estimate of the effect for that category, a code to indicate if the data in that category is undefined (i.e., unclassified) on any of the variables included in the effect, and a label of the variables that describe the data category. If the data in a category is undefined on one of the variables included in the effect, one of the labels used to describe the data will be "UNCLASS". The end of this array is signaled by the number 6 in column 4.

Figure A1. The Mips parameter file of the battle casualty data.

```
*** PARAMETER FILE=W.MIP                      DVNAME=Casualty Count      ***
*** Array 1. (IV #, # OF LEVELS, NAME, NESTING) ***
   1   1  13SHIP        0
   1   2   4WEAPON      0
   1   3   2CASTYPE     1
   5
*** Array 2. (IV #, LEVEL #, LABELS, LIMIT) ***
   2   1   1AM         133
   2   1   2AO         288
   2   1   3APA        679
   2   1   4APD        209
   2   1   5BB        2584
   2   1   6CA        1618
   2   1   7CL        1292
   2   1   8CV        3448
   2   1   9CVE       1080
   2   1  10DD         382
   2   1  11DE         297
   2   1  12DMS        270
   2   1  13LST        181
   2   2   1BOMB         0
   2   2   2GUNFIRE      0
   2   2   3KAMIKAZE     0
   2   2   4TORPEDO      0
   2   3   1WIA          0
   2   3   2KIA          0
   5
*** Array 3. (EFFECT#, COMPONENT#, IV, IV LABEL) ***
   3   1   1   1   1SHIP    :  PROVISIONAL MEAN=    29.680
   3   1   2   0   0
   3   1   3   0   0
   3   2   1   2   1WEAPON  :  PROVISIONAL MEAN=    29.680
   3   2   2   0   0
   3   2   3   0   0
   3   3   1   3   1CASTYPE :  PROVISIONAL MEAN=    29.680
   3   3   2   0   0
   3   3   3   0   0
   3   4   1   1   1SHIP    :  PROVISIONAL MEAN=    29.680
   3   4   2   2  14WEAPON
   3   4   3   0   0
   3   5   1   2   1WEAPON  :  PROVISIONAL MEAN=    29.680
   3   5   2   3   5CASTYPE
   3   5   3   0   0
   5              GRAND MEAN=    29.680
```

34

```
*** Array 4.   (EFFECT#, CELL#, NUM, DEN, EW, EP, CLASS, LABEL) ***
    4  1  1    351.000    44.000   -21.702    0.000  1 AM
    4  1  2    789.000    26.000     0.667    0.000  1 AO
    4  1  3   1247.000    56.000    -7.412    0.000  1 APA
    4  1  4    975.000    52.000   -10.930    0.000  1 APD
    4  1  5   2486.000    82.000     0.638    0.000  1 BB
    4  1  6   3152.000    58.000    24.665    0.000  1 CA
    4  1  7   3320.000    86.000     8.925    0.000  1 CL
    4  1  8   4826.000    76.000    33.820    0.000  1 CV
    4  1  9   2817.000    46.000    31.560    0.000  1 CVE
    4  1 10  11897.000   440.000    -2.641    0.000  1 DD
    4  1 11   1752.000    74.000    -6.004    0.000  1 DE
    4  1 12    631.000    34.000   -11.121    0.000  1 DMS
    4  1 13   1966.000   146.000   -16.214    0.000  1 LST
    4  1 14      0.000     0.000     0.000    0.000  0 UNCLASS
    4  2  1   6539.000   232.000    -1.494    0.000  1 BOMB
    4  2  2   5502.000   314.000   -12.157    0.000  1 GUNFIRE
    4  2  3  13735.000   464.000    -0.078    0.000  1 KAMIKAZE
    4  2  4  10433.000   210.000    20.001    0.000  1 TORPEDO
    4  2  5      0.000     0.000     0.000    0.000  0 UNCLASS
    4  3  1  18661.000   610.000     0.912    0.000  1 WIA
    4  3  2  17548.000   610.000    -0.912    0.000  1 KIA
    4  3  3      0.000     0.000     0.000    0.000  0 UNCLASS
    4  4  1    129.000    18.000     0.684    0.000  1 AM       X BOMB
    4  4  2    521.000    12.000    14.565    0.000  1 AO       X BOMB
    4  4  3    218.000    10.000     1.026    0.000  1 APA      X BOMB
    4  4  4    147.000     6.000     7.244    0.000  1 APD      X BOMB
    4  4  5    191.000     8.000    -4.948    0.000  1 BB       X BOMB
    4  4  6    123.000     8.000   -37.476    0.000  1 CA       X BOMB
    4  4  7   1072.000    24.000     7.556    0.000  1 CL       X BOMB
    4  4  8   1934.000    26.000    12.379    0.000  1 CV       X BOMB
    4  4  9     37.000     4.000   -50.495    0.000  1 CVE      X BOMB
    4  4 10   1709.000    78.000    -3.634    0.000  1 DD       X BOMB
    4  4 11      3.000     2.000   -20.682    0.000  1 DE       X BOMB
    4  4 12     66.000     2.000    15.935    0.000  1 DMS      X BOMB
    4  4 13    389.000    34.000    -0.530    0.000  1 LST      X BOMB
    4  4 14      0.000     0.000     0.000    0.000  0 UNCLASS  X BOMB
    4  4 15     40.000     4.000    14.180    0.000  1 AM       X GUNFIRE
    4  4 16      5.000     2.000   -15.689    0.000  1 AO       X GUNFIRE
    4  4 17     41.000    18.000    -7.833    0.000  1 APA      X GUNFIRE
    4  4 18    187.000     8.000    16.782    0.000  1 APD      X GUNFIRE
    4  4 19    552.000    26.000     3.071    0.000  1 BB       X GUNFIRE
    4  4 20    945.000    20.000     0.062    0.000  1 CA       X GUNFIRE
    4  4 21    218.000    26.000   -18.063    0.000  1 CL       X GUNFIRE
    4  4 22     51.000     4.000   -38.593    0.000  1 CV       X GUNFIRE
    4  4 23    413.000     6.000    19.751    0.000  1 CVE      X GUNFIRE
    4  4 24   2663.000   136.000     4.699    0.000  1 DD       X GUNFIRE
```

4	4	25	253.000	10.000	13.782	0.000	1	DE	X GUNFIRE
4	4	26	17.000	4.000	-2.152	0.000	1	DMS	X GUNFIRE
4	4	27	217.000	50.000	3.031	0.000	1	LST	X GUNFIRE
4	4	28	0.000	0.000	0.000	0.000	0	UNCLASS	X GUNFIRE
4	4	29	79.000	20.000	-3.949	0.000	1	AM	X KAMIKAZE
4	4	30	19.000	4.000	-25.518	0.000	1	AO	X KAMIKAZE
4	4	31	871.000	22.000	17.401	0.000	1	APA	X KAMIKAZE
4	4	32	548.000	36.000	-3.450	0.000	1	APD	X KAMIKAZE
4	4	33	1032.000	36.000	-1.572	0.000	1	BB	X KAMIKAZE
4	4	34	231.000	10.000	-31.167	0.000	1	CA	X KAMIKAZE
4	4	35	650.000	14.000	7.902	0.000	1	CL	X KAMIKAZE
4	4	36	2386.000	34.000	6.755	0.000	1	CV	X KAMIKAZE
4	4	37	1689.000	32.000	-8.380	0.000	1	CVE	X KAMIKAZE
4	4	38	5075.000	162.000	4.367	0.000	1	DD	X KAMIKAZE
4	4	39	504.000	38.000	-10.334	0.000	1	DE	X KAMIKAZE
4	4	40	361.000	26.000	-4.596	0.000	1	DMS	X KAMIKAZE
4	4	41	290.000	30.000	-3.721	0.000	1	LST	X KAMIKAZE
4	4	42	0.000	0.000	0.000	0.000	0	UNCLASS	X KAMIKAZE
4	4	43	103.000	2.000	23.521	0.000	1	AM	X TORPEDO
4	4	44	244.000	8.000	-19.848	0.000	1	AO	X TORPEDO
4	4	45	117.000	6.000	-22.769	0.000	1	APA	X TORPEDO
4	4	46	93.000	2.000	7.749	0.000	1	APD	X TORPEDO
4	4	47	711.000	12.000	° 931	0.000	1	BB	X TORPEDO
4	4	48	1953.000	20.000	...304	0.000	1	CA	X TORPEDO
4	4	49	1380.000	22.000	4.121	0.000	1	CL	X TORPEDO
4	4	50	455.000	12.000	-45.585	0.000	1	CV	X TORPEDO
4	4	51	678.000	4.000	88.259	0.000	1	CVE	X TORPEDO
4	4	52	2450.000	64.000	-8.759	0.000	1	DD	X TORPEDO
4	4	53	992.000	24.000	-2.344	0.000	1	DE	X TORPEDO
4	4	54	187.000	2.000	54.940	0.000	1	DMS	X TORPEDO
4	4	55	1070.000	32.000	-0.030	0.000	1	LST	X TORPEDO
4	4	56	0.000	0.000	0.000	0.000	0	UNCLASS	X TORPEDO
4	4	57	0.000	0.000	0.000	0.000	0	AM	X UNCLASS
4	4	58	0.000	0.000	0.000	0.000	0	AO	X UNCLASS
4	4	59	0.000	0.000	0.000	0.000	0	APA	X UNCLASS
4	4	60	0.000	0.000	0.000	0.000	0	APD	X UNCLASS
4	4	61	0.000	0.000	0.000	0.000	0	BB	X UNCLASS
4	4	62	0.000	0.000	0.000	0.000	0	CA	X UNCLASS
4	4	63	0.000	0.000	0.000	0.000	0	CL	X UNCLASS
4	4	64	0.000	0.000	0.000	0.000	0	CV	X UNCLASS
4	4	65	0.000	0.000	0.000	0.000	0	CVE	X UNCLASS
4	4	66	0.000	0.000	0.000	0.000	0	DD	X UNCLASS
4	4	67	0.000	0.000	0.000	0.000	0	DE	X UNCLASS
4	4	68	0.000	0.000	0.000	0.000	0	DMS	X UNCLASS
4	4	69	0.000	0.000	0.000	0.000	0	LST	X UNCLASS
4	4	70	0.000	0.000	0.000	0.000	0	UNCLASS	X UNCLASS
4	5	1	3121.000	116.000	-2.192	0.000	1	BOMB	X WIA

```
4   5   2   3114.000    157.000     1.400   0.000   1 GUNFIRE  X WIA
4   5   3   8799.000    232.000     7.413   0.000   1 KAMIKAZE X WIA
4   5   4   3627.000    105.000   -16.050   0.000   1 TORPEDO  X WIA
4   5   5      0.000      0.000     0.000   0.000   0 UNCLASS  X WIA
4   5   6   3418.000    116.000     2.192   0.000   1 BOMB     X KIA
4   5   7   2388.000    157.000    -1.400   0.000   1 GUNFIRE  X KIA
4   5   8   4936.000    232.000    -7.413   0.000   1 KAMIKAZE X KIA
4   5   9   6806.000    105.000    16.050   0.000   1 TORPEDO  X KIA
4   5  10      0.000      0.000     0.000   0.000   0 UNCLASS  X KIA
4   5  11      0.000      0.000     0.000   0.000   0 BOMB     X UNCLASS
4   5  12      0.000      0.000     0.000   0.000   0 GUNFIRE  X UNCLASS
4   5  13      0.000      0.000     0.000   0.000   0 KAMIKAZE X UNCLASS
4   5  14      0.000      0.000     0.000   0.000   0 TORPEDO  X UNCLASS
4   5  15      0.000      0.000     0.000   0.000   0 UNCLASS  X UNCLASS
6
```

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION		1b. RESTRICTIVE MARKINGS
Unclassified		N/A

2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT
N/A	Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
N/A	

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
Report No. 91-6	

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
Naval Health Research Center		Chief, Bureau of Medicine and Surgery

6c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code)
P. O. Box 85122	Navy Department
San Diego, CA 92186-5122	Washington, DC 20372-5120

8a. NAME OF FUNDING/SPONSORING ORGANIZATION Naval Medical Research & Development Command	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS			
NNMC Bethesda, MD 20889-5044	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
	63706N	M0095	.005	DN249506

11. TITLE (Include Security Classification)
(U) Using the Medical Information Projection System to Organize and Present Battle Casualty Data

12. PERSONAL AUTHOR(S)
Riemer, Richard A.

13a. TYPE OF REPORT	13b. TIME COVERED		14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM _____	TO _____	1991 April 16	38

16. SUPPLEMENTARY NOTATION

17	COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Casualty estimation, data projections, multiple regression,
			medical planning

19 ABSTRACT (Continue on reverse if necessary and identify by block number)
Medical resource planning requires projections of the number of casualties expected during military operations. Though there is a considerable amount of archival data on casualties, this information is not generally in a format compatible with casualty rate projections. The objective of this study is to show how battle casualty data from WWII can be adapted to the Medical Information Projection system (MIPS). In addition, this document presents the features of MIPS that enhance the manipulation and evaluation of data. The statistical procedures of MIPS are based on the general linear model as it applies to multiple regression and the analysis of variance. These procedures are used to calculate the magnitude of effect for the predictor variables included in the statistical model. However, statistical significance is not calculated because MIPS is designed to be used with large archival data sets that may not meet the parametric characteristics needed for accurate significance testing. After MIPS calculates the magnitude of effects for the predictor variables the user may query MIPS for data projections, in this instance, casualty estimates.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	Unclassified

22a NAME OF RESPONSIBLE INDIVIDUAL	22b TELEPHONE (Include Area Code)	22c OFFICE SYMBOL
William Pugh	553-8403	20